signal being outputted by the audio encoder **921'** as first encoded representation **921** and second encoded representation **922** for each subband of the at least one subband of the plurality of subbands. Furthermore, the audio encoder **920'** outputs an audio codec indicator **925** being indicative that the at least one subband of the plurality of subbands is encoded in accordance with the selected first audio codec. For instance, the at least one subband may for instance be a subset of subbands of the plurality of subbands or all subbands of the plurality of subbands.

[0447] Accordingly, the term "bypass" has to be understood in a way that the first encoded signal representation **921** and the second encoded signal representation **922** outputted by the audio encoder **910**, **910'** does not depend or is not influenced by the combining operation of the first audio coded, e.g. as performed by the combining entity **941**.

[0448] Thus, as an example, the first and second signal representation may be bypassed with respect to the combining operation of the first audio codec in a way that the first signal representation is outputted by the audio decoder **920'** as the first encoded representation and the second signal representation is outputted by the audio decoder **921'** as the second encoded representation.

[0449] FIG. **10** depicts a schematic block diagram of a second example embodiment of an apparatus **1000** according to the third aspect of invention.

[0450] For instance, this apparatus **1000** may be based on the apparatus **910** depicted in FIG. **9**. The apparatus **1000** comprises an audio encoder **1020**, which may represent the audio encoder **920** depicted in FIG. **9***a* or the audio encoder **920'** depicted in FIG. **9***c*.

[0451] In FIG. **10**, the first signal representation is indicated by reference sign **1001** and the second signal representation is indicated by reference sign **1002**.

[0452] If the first and second signal representation **1001**, **1002** are not in the frequency-domain, i.e., if the first and the second signal representation are in the time domain then the first signal representation **1001** is fed to an optional entity for block division and windowing **1011**, wherein this entity **1011** may be configured to generate windows with a predefined overlap and an effective length, wherein this predefined overlap map represent 50 or another well-suited percentage, and wherein this effective length may be 20 ms or another well-suited length.

[0453] Furthermore, the entity **1011** may be configured to add $D_{tot}=D_{max}+D_{HRTF}$ zeroes to the end of the window, wherein D_{max} may correspond to the maximum delay in samples between the microphones, as explained with respect to the method depicted in FIG. **3**.

[0454] Similarly, the optional entity for block division and windowing **1012** may receive the second signal representation and is configured to generate windows with a predefined overlap and an effective length in the same way as optional entity **1011**.

[0455] The windows formed by entities configured to generate windows with a predefined overlap and an effective length **1011**, **1012** are fed to the respective optional transform entity **1021**, **1022**, wherein transform entity **1021** is configured to transform the windows of the first signal representation **1001** to frequency domain, and wherein transform entity **1022** is configured to transform the windows of the second signal representation **1002** to frequency domain. This may be done in accordance with the explanation presented with respect to step **320** of FIG. **3***a*.

[0456] Thus, transform entity **421** may be configured to output $S_1(n)$ and transform entity **422** may be configured to output $S_2(n)$.

[0457] If the first and second signal representation **1001**, **1002** are in the frequency-domain, then optional entities **1011**, **1012**, **1021** and **1022** may be omitted and the first signal representation **1001** can be used as first signal representation **931** which is fed as input signal **911** to the audio encoder **1020** and the second signal representation **1002** can be used as second signal representation **932** which is fed to the audio encoder **1020**.

[0458] The audio encoder **1020** outputs the first encoded signal representation **921** and the second encoded signal representation **922**, as explained above. Furthermore, the audio encoder **1020** outputs an audio codec indicator **925** being indicative that the at least one subband of the plurality of subbands is encoded in accordance with the selected first audio codec, as explained above.

[0459] Entity **1030** is configured to perform quantization end encoding to the first encoded signal representation $A_1(n)$ in the frequency domain and to the second encoded signal representation $A_2(n)$ in the frequency domain For instance, suitable audio codes may for instance be AMR-WB+, MP3, AAC and AAC+, or any other audio codec.

[0460] Afterwards, the quantized and encoded first and second signal representations **1031**, **1032** are inserted into a bitstream **1050** by means of bitstream generation entity **1040**.

[0461] The directional information **935** associated with at least one subband of the plurality of subbands associated with the left and the right signal representation is inserted into the bitstream **1005** by means of the bitstream generation entity **440**. Furthermore, for instance, the directional information **403** may be quantized and/or encoded before being inserted in the bitstream **1005**. This may be performed by entity **1030** (not depicted in FIG. **10**).

[0462] Thus, the apparatus **1000** is configured to output an encoded audio representation **1050** comprising the first and second signal representation **1001**, **1002**, the directional information **935**, and the indicator **935**.

[0463] As will be exemplarily described with respect to the apparatus **1100** depicted in FIG. **11**, the encoded audio representation **1050** might be considered to represent a backward compatible audio representation which may be encoded to the left and right signals by an audio decoder which is configured to perform audio decoding according to the first audio codec.

[0464] Apparatus **1100** comprises an audio decoder **1120**, which is configured to receive a first encoded signal representation **1116** and a second signal representation **1117** and which is configured to perform an audio decoding in accordance with the first audio codec for each subband which is indicated to be encoded with the first audio coded by the indicator **1111**.

[0465] The apparatus **1100** receives an encoded audio representation **1101**, which may represent or be based on the encoded audio representation **1050** depicted in FIG. **10**.

[0466] A bitstream entity **1110** is configured to extract the indicator from the encoded audio representation **1101**, which is fed as indicator **1111** to the audio decoder **1120**. Furthermore, the bitstream entity feeds the encoded first and second signal representation **1112**, **1113** to an entity for decoding and inverse quantization **1115**. This entity for decoding and inverse quantization **1115** may represent the counterpart to the entity for quantization and coding **1030** depicted in FIG. **10**, i.e. the entity for decoding and inverse quantization **1115**